

Multiscale System Theory

Albert Benveniste
IRISA-INRIA, Campus de Beaulieu
35042 RENNES CEDEX, FRANCE

Ramine Nikoukhah
INRIA, Rocquencourt, BP 105
78153 LE CHESNAY CEDEX, FRANCE

Alan S. Willsky *
LIDS, MIT
CAMBRIDGE, MA 02139, USA

February 21, 1990

*The work of this author was supported in part by the Air Force Office of Scientific Research under grant AFOSR-88-0032, in part by the National Science Foundation under Grant ECS-3700903 and in part by the US Army Research Office under Contract DAAL03-86-K-0171.

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 21 FEB 1990		2. REPORT TYPE		3. DATES COVERED 00-02-1990 to 00-02-1990	
4. TITLE AND SUBTITLE Multiscale System Theory				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Massachusetts Institute of Technology,Laboratory for Information and Decision Systems,77 Massachusetts Avenue,Cambridge,MA,02139-4307				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 29	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Abstract

In many applications, it is of interest to analyze and recognize phenomena occurring at different scales. The recently introduced wavelet transforms provide a time-and-scale decomposition of signals that offers the possibility of such an analysis. Until recently, however, there has been no corresponding statistical framework to support the development of optimal, multiscale statistical signal processing algorithms. A recent work of some of the present authors and co-authors proposed such a framework via models of “stochastic fractals” on the dyadic tree. In this paper we investigate some of the fundamental issues that are relevant to system theories on the dyadic tree, both for systems and signals.

1 Introduction

The investigation of multi-scale representations of signals and the development of multiscale algorithms have been and remain topics of much interest in many contexts. In some cases, such as in the use of fractal models for signals and images [4, 30] the motivation has directly been the fact that the phenomenon of interest exhibits patterns of importance at multiple scales. A second motivation has been the possibility of developing highly parallel and iterative algorithms based on such representations. Multigrid methods for solving partial differential equations [9, 24, 31, 33] or for performing Monte Carlo experiments [14] are good examples. A third motivation stems from the so-called “sensor fusion” problems in which one is interested in combining measurements with very different spatial resolutions. Geophysical problems, for example, often have this character. Finally, renormalization group ideas, from statistical physics, now find application in methods for improving convergence in large-scale simulated annealing algorithms for Markov random field estimation [19].

One of the more recent areas of investigation in multi-scale analysis has been the development of a theory of multi-scale representations of signals [27, 29] and the closely related topic of wavelet transforms [15, 21, 25, 16, 23, 17, 22]. These methods have drawn considerable attention in several disciplines including signal processing because they appear to be a natural way to perform a time-scale decomposition of signals and because examples that have been given of such transforms seem to indicate that it should be possible to develop efficient optimal processing algorithms based on these representations. The development of such optimal algorithms—e.g. for the reconstruction of noise-degraded signals or for the detection and localization of transient signals of different duration—requires, of course, the development of a corresponding system theory and a theory of stochastic processes and their estimation. The research presented in this and several other papers and reports [13, 7, 14, 5] has the development of this theory as its objective.

In the next section, we introduce multi-scale representations of signals and wavelet transforms and from these we motivate the investigation of models on dyadic trees. Then we report some facts on the geometry of dyadic trees that are essential in our theory. In particular we introduce $1/$ shift operators to encode any move on the tree, and $2/$ translations on the tree, and we show that these are entirely different notions, as opposed to the case of classical $1D$ - and $2D$ - systems. In section 3 we develop a sample of a system theory of “general”

transfer functions, i.e. a theory where transfer functions are considered as formal power series in the primitive shift operators. We show that this theory is a particular case of *noncommutative formal power series* theory introduced and studied mainly by M. Fliess [8]. This theory however does not seem to be a reasonable basis for a theory of stochastic processes on the dyadic tree. Such a topic is the purpose of Section 4. In this section, *stationary* transfer functions are defined as transfer functions commuting with any translation. We characterize in a simple way such stationary transfer functions and develop a realization theory for them, showing by the way that such a theory relates to S. Attasi's *2D*-system theory [3]. Then we define *stationary stochastic processes* as stochastic processes with translation invariant covariance function, and we present a sample of first basic properties of such processes.

In [5, 7, 6], we have studied extensively a subclass of the stationary processes, namely the class of *isotropic processes*, i.e. processes with isometry invariant covariance functions; we have also developed Schur- and Levinson-like parametrizations for such processes. However, we were not able to develop a corresponding theory of "isotropic" transfer functions, and the present paper fills this gap by providing a clean system theory for a more general class of systems.

Due to lack of space, results are stated without proof in this paper, proofs will be presented in a full paper in preparation.

2 Multiscale Representations and Stochastic Processes on Homogeneous Trees

2.1 Multiscale Representations and Wavelet Transforms

The multi-scale representation [28, 29] of a continuous signal $f(x)$ consists of a sequence of approximations of that signal at finer and finer scales where the approximation of $f(x)$ at the m th scale is given by

$$f_m(x) = \sum_{n=-\infty}^{+\infty} f(m, n)\phi(2^m x - n) \quad (2.1)$$

As $m \rightarrow \infty$ the approximation consists of a sum of many highly compressed, weighted, and shifted versions of the function $\phi(x)$ whose choice is far from arbitrary. In particular in order for the $(m + 1)$ st approximation to be a refinement of the m th, we require that $\phi(x)$ be exactly representable at the next scale:

$$\phi(x) = \sum_n h(n)\phi(2x - n) \quad (2.2)$$

Furthermore in order for (2.1) to be an orthogonal series, $\phi(t)$ and its integer translates must form an orthogonal set. As shown in [16], $h(n)$ must satisfy several conditions for this and several other properties of the representation to hold. In particular $h(n)$ must be the impulse response of a quadrature mirror filter [16, 34]. The simplest example of such a ϕ, h pair is the Haar approximation with

$$\phi(x) = \begin{cases} 1 & 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (2.3)$$

and

$$h(n) = \begin{cases} 1 & n = 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.4)$$

Multiscale representations are closely related to wavelet transforms. Such a transform is based on a single function $\psi(x)$ that has the property that the full set of its scaled translates $\{2^{m/2}\psi(2^m x - n)\}$ form a complete orthonormal basis for L^2 . In [16] it is shown that ϕ and ψ are related via an equation of the form

$$\psi(x) = \sum_n g(n)\phi(2x - n) \quad (2.5)$$

where $g(n)$ and $h(n)$ form a *conjugate mirror filter* pair [34], and that

$$f_{m+1}(x) = f_m(x) + \sum_n d(m, n)\psi(2^m x - n) \quad (2.6)$$

$f_m(x)$ is simply the partial orthonormal expansion of $f(x)$, up to scale m , with respect to the basis defined by ψ . For example if ϕ and h are as in eq. (2.3), eq. (2.4), then

$$\psi(x) = \begin{cases} 1 & 0 \leq x < 1/2 \\ -1 & 1/2 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (2.7)$$

$$g(n) = \begin{cases} 1 & n = 0 \\ -1 & n = 1 \\ 0 & \text{otherwise} \end{cases} \quad (2.8)$$

and $\{2^{m/2}\psi(2^m x - n)\}$ is the *Haar basis*.

From the preceding remarks we see that we have a *dynamical* relationship between the coefficients $f(m, n)$ at one scale and those at the next scale. Indeed this relationship defines a lattice on the points (m, n) , where $(m + 1, k)$ is connected to (m, n) if $f(m, n)$ influences $f(m + 1, k)$. In particular the Haar representation naturally defines a dyadic tree structure on the points (m, n) in which each point has two descendents corresponding to the two subdivisions of the support interval of $\phi(2^m x - n)$, namely those of $\phi(2^{(m+1)}x - 2n)$ and $\phi(2^{(m+1)}x - 2n - 1)$. This observation provides the motivation for the development of models for stochastic processes on dyadic trees and associated system theory as the basis for a statistical theory of multiresolution stochastic processes.

Let us discuss briefly how particular such a system theory may be. In classical system theory on \mathbf{Z} , an important tool is the z -transform. In this case, the shift operator z is used both for defining what *stationary* means for a linear operator on sequences (commuting with z , viewed as a primitive translation), and for *encoding* transfer functions. Moreover, \mathbf{Z} is totally ordered so that there is an obvious notion of causality which plays an important role in system theory. In the $2D$ -case of time index set \mathbf{Z}^2 , the natural definition of causality is lost but other features remain. As we shall see throughout this paper, the situation is drastically different for system theory on homogeneous trees. Although less natural than for \mathbf{Z} , causality may be reasonably introduced far less arbitrarily than for the $2D$ case. However natural shift operators that encode “transfer functions” in our case will *not* be translations, not even isometries; on the other hand, translations may be defined as we shall see next, but they cannot be represented using the shift operators. This strange situation has deep consequences which we shall investigate throughout this paper.

2.2 Homogeneous Trees

Homogeneous trees, and their structure, have been the subject of some work [1, 2, 12, 18, 11] in the past on which we build and which we now briefly review. A *homogeneous tree* \mathcal{T} of order q is an infinite acyclic, undirected, connected graph such that every node of \mathcal{T} has exactly $(q + 1)$ branches to other nodes. Note that $q = 1$ corresponds to the usual integers with the obvious branches from one integer to its two neighbors. The case of $q = 2$, illustrated in Figure 1, corresponds, as we will see, to the dyadic tree on which we focus in this paper. In 2-D signal processing, it would be natural to consider the case of $q = 4$ leading to a pyramidal structure on the indexing set of the 2-D processes.

Isometries. The tree \mathcal{T} has a natural notion of distance: $d(s, t)$ is the number of branches along the shortest path between the nodes $s, t \in \mathcal{T}$ (by abuse of notation we use \mathcal{T} to denote both the tree and its collection of nodes). One can then define the notion of an isometry on \mathcal{T} which is simply a one-to-one map of \mathcal{T} onto itself that preserves distance. For the case of $q = 1$, the group of all possible isometries corresponds to translations of the integers ($t \mapsto t + k$), the reflection operation ($t \mapsto -t$), and concatenations of the two. For $q \geq 2$ the group of isometries of \mathcal{T} is significantly larger and more complex. The following classification of isometries may be found in [12]:

Lemma 1 (classification of isometries) *Given an isometry f of the homogeneous tree \mathcal{T} , three cases are possible, namely:*

$$\exists s \in \mathcal{T} : f(s) = s \quad (2.9)$$

$$\exists s, t \in \mathcal{T} : d(s, t) = 1 \text{ and } f(s) = t, f(t) = s \quad (2.10)$$

$$\exists (s_n)_{n \in \mathbb{Z}} \in \mathcal{T}, \exists i > 0 : d(s_n, s_{n+1}) = 1, f(s_n) = s_{n+i} \quad (2.11)$$

Boundary points and horocycles. An important concept here is the notion of a *boundary point* [2, 11] of a tree. Consider the set of infinite sequences of \mathcal{T} where any such sequence consists of a sequence of distinct nodes t_1, t_2, \dots where $d(t_i, t_{i+1}) = 1$. A boundary point is an equivalence class of such sequences where two sequences are equivalent if they differ by a finite number of nodes. For $q = 1$, there are only two such boundary points corresponding to sequences increasing towards $+\infty$ or decreasing towards $-\infty$. For $q = 2$ the set of boundary points is uncountable. In this case let us choose one boundary point which we denote by $-\infty$.

Once we have distinguished this boundary point, we can identify a partial order on \mathcal{T} . In particular note that from any node t there is a unique path in the equivalence class defined by $-\infty$ (i.e. a unique path from t “towards” $-\infty$). Then if we take any two nodes s and t , their paths to $-\infty$ must differ only by a finite number of points and thus must meet at some node which we denote by $s \wedge t$ (see Figure 1). Thus, we can define a notion of *relative distance* of two nodes to $-\infty$:

$$\delta(s, t) = d(s, s \wedge t) - d(t, s \wedge t) \quad (2.12)$$

so that

$$s \preceq t \text{ (“} s \text{ is at least as close to } -\infty \text{ as } t \text{”) if } \delta(s, t) \leq 0 \quad (2.13)$$

$$s \prec t \text{ (“} s \text{ is closer to } -\infty \text{ than } t \text{”) if } \delta(s, t) < 0 \quad (2.14)$$

This also yields an equivalence relation on nodes of \mathcal{T} :

$$s \asymp t \leftrightarrow \delta(s, t) = 0 \quad (2.15)$$

For example, the points s , v , and u in Figure 1 are all equivalent. The equivalence classes of such nodes are referred to as *horocycles*.

These equivalence classes can best be visualized as in Figure 2 by redrawing the tree, in essence by picking the tree up at $-\infty$ and letting the tree “hang” from this boundary point. In this case the horocycles appear as points on the same horizontal level and $s \preceq t$ means that s lies on a horizontal level above or at the level of t . Note that in this way we make explicit the dyadic structure of the tree. With regard to multiscale signal representations, a shift on the tree toward $-\infty$ corresponds to a shift from a finer to a coarser scale and points on the same horocycle correspond to the points at different translational shifts in the signal representation at a single scale.

Translations. Translations will play an important role in the definition of stationarity. Translations certainly should be isometries of the third class (cf. (2.11)) according to lemma 1. However, for the sequel, we shall need *primitive translations* encoding “moving away from $-\infty$ ”, i.e. the counterpart of the shift operator z on \mathbb{Z} . These are defined as follows:

1. select an infinite path $(t_n)_{n \in \mathbb{Z}}$ originating from $-\infty$, call it the *skeleton* of the translation,
2. denote by s_n the unique point outside the skeleton such that $d(s_n, t_n) = 1$

3. denote by $\mathcal{T}_{s_n}^+$ the semi infinite dyadic tree with root s_n composed of the semi infinite paths originating at s_n and moving away from $-\infty$
4. then the *translation with skeleton* (t_n) is the unique isometry τ such that (cf. Figure 3)

$$\tau(t_n) = t_{n+1}, \quad \tau(\mathcal{T}_{s_n}^+) = \mathcal{T}_{s_{n+1}}^+ \quad (2.16)$$

2.3 Shift operators and transfer functions on \mathcal{T}

Shifts on the tree. The counterpart of the shift operator z is composed of the *two* shifts which are illustrated in Figure 2

- 1 the identity operator (no move)
- α the left down-shift (move one step away from $-\infty$ toward the left)
- β the right down-shift (move one step away from $-\infty$ toward the right)

These shifts act on the right (if t is any node on the tree, $t\alpha$ is its left offspring). Note that α and β are one-to-one but not onto; they are *not* isometries.

Shift operators on signals. By “signal” we mean a family y_t of scalars or vectors indexed by the vertices of the tree. The primitive operators that we consider are “dual” of the shifts on \mathcal{T} , namely (see figure 2):

- 1 the identity operator (no move)
- α the left down-shift operator:¹

$$y = \alpha u \Leftrightarrow \forall t : y_t = u_{t\alpha}$$

- β the right down-shift operator:

$$y = \beta u \Leftrightarrow \forall t : y_t = u_{t\beta}$$

- $\bar{\alpha}$ the right up-shift operator:²

$$y = \bar{\alpha} u \Leftrightarrow \forall t : \begin{cases} y_{t\alpha} = u_t \\ y_{t\beta} = 0 \end{cases}$$

¹the value of y at a given node is obtained by picking the value of u at the corresponding *left down* node

²the value of y at a given node is obtained by picking the value of u at the corresponding *right up* node if available, or by setting 0 otherwise

- $\bar{\beta}$ the left up-shift operator:

$$y = \bar{\beta}u \Leftrightarrow \forall t : \begin{cases} y_{t\beta} = u_t \\ y_{t\alpha} = 0 \end{cases}$$

The class of operators we consider is the linear space over \mathbf{R} spanned by these primitive shifts: this is a noncommutative algebra. We shall call *transfer functions* the matrices the entries of which are elements of this algebra. It is easy to verify that the primitive operators obey the following simplification rules:

$$\alpha\bar{\alpha} = \beta\bar{\beta} = 1 \quad (2.17)$$

$$\beta\bar{\alpha} = \alpha\bar{\beta} = 0 \quad (2.18)$$

$$\bar{\alpha}\alpha + \bar{\beta}\beta = 1 \quad (2.19)$$

Thanks to these rules, any transfer function may be expressed as follows:

$$S = \sum_{\substack{w^\dagger \in \mathcal{W}^\dagger \\ w^\downarrow \in \mathcal{W}^\downarrow}} s_{w^\dagger w^\downarrow} w^\dagger w^\downarrow \quad (2.20)$$

where \mathcal{W}^\dagger and \mathcal{W}^\downarrow are the family of monomials generated by the up-shifts $\bar{\alpha}, \bar{\beta}$ and the down-shifts α, β respectively, and the $s_{w^\dagger w^\downarrow}$'s are matrix coefficients. In this writing we implicitly assume that all simplifications (2.17, 2.18, 2.19) have been performed. This means that any monomial may be decomposed into a down-shift followed by an up-shift.

We shall call the *support* of S the set of monomials in (2.20) with nonzero coefficient.

Causality. We shall say that a monomial $w^\dagger w^\downarrow$ is *causal* if

$$\text{degree}(w^\dagger) \geq \text{degree}(w^\downarrow) \quad (2.21)$$

and we say that the transfer function S is *causal* if, in expression (2.20), $s_{w^\dagger w^\downarrow} = 0$ whenever $w^\dagger w^\downarrow$ is non-causal. Strict causality is defined accordingly. Causal transfer functions may be written as follows

$$S = \sum_{\substack{w^\dagger \in \mathcal{W}^\dagger \\ \vec{w} \in \vec{\mathcal{W}}}} s_{w^\dagger \vec{w}} w^\dagger \vec{w} \quad (2.22)$$

where $\vec{\mathcal{W}}$ is the set of monomials $w^\dagger w^\downarrow$ such that

$$\text{degree}(w^\dagger) = \text{degree}(w^\downarrow)$$

i.e. monomials in $\vec{\mathcal{W}}$ combine data from the considered horocycle.

3 System theory and realization of general causal transfer functions

In this section, we investigate some aspects of system theory for the notion of transfer function introduced in the preceding section. Here we consider “general” (not necessarily stationary) transfer functions; stationary systems will be studied in Section 3. We shall see that the theory of general transfer functions is related to realization theory for automata [8] rather than linear system theory even though we are considering linear operators on signals.

Definition 1 We define the depth of a causal monomial $w = w^\dagger \vec{w}$ (cf. formula (2.22)) as one half the degree of \vec{w} . A transfer function S is called finite depth if it can be expressed as a sum of bounded depth monomials.

The following lemma is obvious:

Lemma 2 If S is finite depth we can decompose it as follows

$$S = S^\dagger \vec{S} \quad (3.1)$$

where S^\dagger is a transfer function with support in \mathcal{W}^\dagger and \vec{S} a finite degree transfer function with support in $\vec{\mathcal{W}}$.

S^\dagger performs a smoothing along the infinite path linking the current point to $-\infty$ while \vec{S} performs a smoothing along the horocycle. Hence the support of a finite depth transfer function is a cylinder as shown in the figure 4.

3.1 State-space realizations

Definition 2 A transfer function S is realizable if there exist constant matrices C, A_α, A_β and a transfer function \vec{S} as in (3.1) such that

$$S = C (I - \bar{\alpha} A_\alpha - \bar{\beta} A_\beta)^{-1} \vec{S} \quad (3.2)$$

A state-space realization of (3.2) is

$$\begin{cases} x &= \bar{\alpha} A_\alpha x + \bar{\beta} A_\beta x + \vec{S} u \\ y &= Cx \end{cases}$$

which is equivalent to

$$\begin{cases} x_{t\alpha} &= A_\alpha x_t + \alpha \vec{S} u_t \\ x_{t\beta} &= A_\beta x_t + \beta \vec{S} u_t \\ y_t &= Cx_t \end{cases}$$

3.2 Realization in the zero depth case

According to (3.1), a zero depth transfer function may be expressed as

$$S = \sum s_w w^\dagger$$

As usually done in automata and noncommutative formal power series theories, we associate with S the following *Hankel matrix*:

$$\mathcal{H}(S)_{ij} = s_{w_i^\dagger w_j^\dagger}$$

where the monomials $(w_i^\dagger)_{i \geq 0}$ are ordered according to the increasing degree with priority given to $\bar{\alpha}$. Then the following results may be borrowed from noncommutative formal power series theory [8]:

Theorem 1 *S is realizable if and only if $\mathcal{H}(S)$ has finite rank. Moreover, the dimension of minimal realizations equals this rank, i.e.*

$$S = C \left(I - \bar{\alpha} A_\alpha - \bar{\beta} A_\beta \right)^{-1} B$$

where the dimensions of A_α and A_β equals the rank of $\mathcal{H}(S)$.

By writing

$$A_w = \text{coefficient of } w \text{ in } \left(I - \bar{\alpha} A_\alpha - \bar{\beta} A_\beta \right)^{-1}$$

we also have:

Theorem 2 *A realization $(C, A_\alpha, A_\beta, B)$ is minimal if and only if*

$$\begin{aligned} \bigvee_{|w| < n} \text{Im}(A_w B) &= \mathbf{R}^n \\ \bigcap_{|w| < n} \text{Ker}(C A_w) &= \{0\} \end{aligned}$$

where n is the dimension of the state and where $|w|$ denotes the total degree of w .

As a corollary, we know that all minimal realizations are related by similarity transformations.

To conclude, nothing new really appears, except that our theory relates to noncommutative power series theory rather than usual $1D$ - or $2D$ -system theories.

3.3 Realization in the k -depth case

The above procedure has to be modified for this case. Consider the space $\vec{\mathcal{W}}_k$ spanned by the monomials \vec{w} of degree $\leq 2k$. Recall that $\bar{\alpha}\alpha + \bar{\beta}\beta = 1$ so that the family of these monomials is not a basis of $\vec{\mathcal{W}}_k$. However, it is easily checked that monomials with degree **exactly** equal to $2k$ form a basis for $\vec{\mathcal{W}}_k$. Denote by $\{\phi_1, \dots, \phi_{n_k}\}$ such a basis, and set

$$\Phi_k = \begin{bmatrix} \phi_1 \\ \cdot \\ \cdot \\ \cdot \\ \phi_{n_k} \end{bmatrix} \otimes I$$

where \otimes denotes the Kronecker product, and the identity matrix I is of suitable dimension for eqn. (3.3) to be consistent. Then we have the following theorem:

Theorem 3 1. *If S has depth k , it can be expressed as*

$$S = S^\dagger \Phi_k \tag{3.3}$$

2. *S is realizable if and only if S^\dagger in (3.3) is realizable.*

3. *If $(C, A_\alpha, A_\beta, B)$ is a minimal realization of S^\dagger then $(C, A_\alpha, A_\beta, B\Phi_k)$ is a minimal realization of S .*

The realization procedure for the k -depth case is:

1. Express S as

$$S = S^\dagger \Phi_k$$

making sure that no simplification is possible.

2. Realize S^\dagger as

$$S^\dagger = C (I - \bar{\alpha}A_\alpha - \bar{\beta}A_\beta)^{-1} B$$

3. Construct the minimal realizations:

$$\begin{cases} x_{t\alpha} &= A_\alpha x_t + \alpha B \Phi_k u_t \\ x_{t\beta} &= A_\beta x_t + \beta B \Phi_k u_t \\ y_t &= C x_t \end{cases}$$

3.4 Discussion

In this section, we have developed a realization theory for finite depth transfer functions. To extend such a theory to infinite depth transfer functions, we need a notion of rationality for the matrix \tilde{S} . This does not seem to exist in general.

On the other hand, the notion of transfer function that we have introduced cannot be considered as “stationary” in any reasonable sense. For instance the relation $y = \bar{\alpha}u$ where $u \equiv 1$ yields $y_{t\alpha} = 1$ but $y_{t\beta} = 0$! This means that, to develop a theory of stationary processes, we need to constrain the class of transfer functions that we have considered so far. This will be the subject of the next section.

4 Stationary causal and noncausal transfer functions and stochastic processes

Given a translation τ of \mathcal{T} , by abuse of notation, we also denote by τ its action on signals defined by

$$\tau(y)_t = y_{\tau(t)}$$

Definition 3 (stationary transfer functions) *A transfer function S is said to be stationary if³*

$$S \circ \tau = \tau \circ S$$

for any primitive translation τ .

To further study this notion, we need to know more about the primitive translations.

4.1 More on primitive translations

It will be convenient to re-encode definition (2.16) of primitive translations using the shift operators on \mathcal{T} . Let $\Gamma = \{t_n\}_{n \in \mathbb{Z}}$ be the skeleton of the considered primitive translation denoted by τ_Γ , and denote by s_n the unique point outside the skeleton such that $d(t_n, s_n) = 1$. Then τ_Γ is encoded by the following formulae

$$\begin{aligned} \tau_\Gamma(t_n) &= t_{n+1} \\ \tau_\Gamma(s_n w^\downarrow) &= s_{n+1} w^\downarrow \end{aligned} \tag{4.1}$$

Given two skeletons Γ and Γ' , we define their *composition*

$$\Gamma'' \triangleq \Gamma \circ \Gamma'$$

by the following formulae, where we label the two skeletons in such a way that they exactly bifurcate after t_0 , i.e. $t_0 = t'_0, t_1 \neq t'_1$ and n denotes an arbitrary nonnegative integer:

$$\begin{aligned} t''_{-n} &= t_{-n} \\ t''_1 &= t_1 \\ t''_{2+n} &= s_1 w^\downarrow \text{ if } t'_{1+n} = t'_1 w^\downarrow \end{aligned} \tag{4.2}$$

³ \circ denotes the composition of maps.

We have the following result:

$$\tau_\Gamma \circ \tau_{\Gamma'} = \tau_{\Gamma \circ \Gamma'}^2 \quad (4.3)$$

A nice consequence of formula (4.3) is that *the family of powers of primitive translations is a semi-group*.

4.2 Characterization of stationary transfer functions

Noncausal transfer functions. Using formulae (4.1,4.3) the following fundamental result may be proved:

Theorem 4 *The transfer function S is stationary if and only if it can be written as follows*

$$S = \sum_{\substack{w^\uparrow \in \mathcal{W}^\uparrow \\ w^\downarrow \in \mathcal{W}^\downarrow}} s_{|w^\uparrow||w^\downarrow|} w^\uparrow w^\downarrow \quad (4.4)$$

Let us introduce the following stationary primitive transfer functions:

$$\gamma = \frac{1}{2}(\alpha + \beta) \quad (4.5)$$

$$\bar{\gamma} = \bar{\alpha} + \bar{\beta} \quad (4.6)$$

These two operators generate two semi-groups. The action of these semi-groups is depicted in Figure 5: $\bar{\gamma}$ is a “backward” shift towards $-\infty$ whereas γ is a “forward-and-average” shift (the “Haar smoother”). Using these operators, formula (4.4) may be rewritten as

$$S = \sum_{k,l \geq 0} s_{k,l} \bar{\gamma}^k \gamma^l \quad (4.7)$$

Causal transfer functions. The γ and $\bar{\gamma}$ operators obey the following simplification rule

$$\gamma \bar{\gamma} = 1 \quad (4.8)$$

It will be useful to introduce the following family of operators which perform a smoothing of data on the same horocycle as shown in the figure 5:

$$\delta^{[k]} = \bar{\gamma}^k \gamma^k \quad (4.9)$$

All $\delta^{[k]}$'s are *idempotent* operators. These operators may be used to provide the following counterpart of formula (2.22) for the stationary case:

Theorem 5 *If S is stationary and causal, it can be expressed as follows:*

$$S = \sum_{k,l \geq 0} s_{k,l} \bar{\gamma}^k \delta^{[l]} \quad (4.10)$$

Obviously the matrix coefficients $s_{k,l}$ are different in formulae (4.7) and (4.10).

4.3 Realization of stationary transfer functions

Both formulae (4.7) and (4.10) may be interpreted as standard $2D$ -transfer functions that are *causal in the two variables*. Hence standard $2D$ realization theories may be applied to both cases. We shall briefly investigate the two cases.

Non causal transfer functions. If we interpret γ as the row operator and $\bar{\gamma}$ as the column operator, then it is natural to consider the row-by-row scanning to define a total ordering on the $2D$ index space. This corresponds to decomposing the transfer function S according to the following two steps:

1. a bottom-up (i.e. fine-to-coarse) smoothing, followed by
2. a top-down (i.e. coarse-to-fine) propagation.

$2D$ -system theory for systems having separable denominator [3] may be applied here. Rational transfer functions in this latter case are of the following form [26]

$$S = C (I - \bar{\gamma} A \bar{\gamma})^{-1} P (I - \gamma A \gamma)^{-1} B \quad (4.11)$$

which yields the following state space form

$$\begin{cases} v_t &= A_\gamma \left(\frac{v_{t\alpha} + v_{t\beta}}{2} \right) + B u_t \\ z_t &= P_2 v_t \\ x_{t\alpha} &= A_{\bar{\gamma}} x_t + P_1 z_{t\alpha} \\ x_{t\beta} &= A_{\bar{\gamma}} x_t + P_1 z_{t\beta} \\ y_t &= C x_t \end{cases} \quad (4.12)$$

where $P = P_1 P_2$. The first two equations define a purely “anticausal” process, whereas the last third equations define a causal zero depth process.

Causal transfer functions. Here we interpret the sequence $\delta^{[k]}$ as the powers of the row operator and $\bar{\gamma}$ as the column operator. Then again we consider the row-by-row scanning to define a total ordering of the $2D$ index space. This corresponds to decomposing the transfer function S according to the following two steps:

1. a smoothing along the considered horocycle (i.e. constant scale smoothing), followed by
2. a top-down (i.e. coarse-to-fine) propagation.

$2D$ -system theory for systems having separable denominator [3] may again be applied here. Rational transfer functions in this latter case are of the following form [26]

$$S = C (I - \bar{\gamma} A_{\bar{\gamma}})^{-1} P (I - \delta A_{\delta})^{-1} B \quad (4.13)$$

where it is understood that, in expanding such a formula into a power series, δ^k should be replaced by $\delta^{[k]}$. This latter unusual feature has for consequence that no tractable time domain translation of the “frequency domain” formula (4.13) is available. The *finite depth* case however yields

$$\begin{cases} x_{t\alpha} &= A_{\bar{\gamma}} x_t + B \left(1, \delta, \dots, \delta^{[k]} \right) u_{t\alpha} \\ x_{t\beta} &= A_{\bar{\gamma}} x_t + B \left(1, \delta, \dots, \delta^{[k]} \right) u_{t\beta} \\ y_t &= C x_t \end{cases} \quad (4.14)$$

where $B \left(1, \delta, \dots, \delta^{[k]} \right)$ is a linear combination of the listed operators. This corresponds to the case where A_{δ} is nilpotent.

It can be shown that stationary finite depth scalar transfer functions may be equivalently expressed in the following *ARMA* form

$$S = A^{-1} B \quad (4.15)$$

where A is a causal transfer function of *finite support* and $B = B \left(1, \delta, \dots, \delta^{[k]} \right)$ is as in (4.14). This ARMA form includes as a special case the AR modeling filters for “isotropic” processes studied in [7, 6, 5].

4.4 Stationary stochastic processes

To simplify the presentation, we concentrate here on scalar processes.

Definition 4 A zero mean stochastic process y is said to be stationary if its covariance function is translation-invariant, i.e.

$$\mathbf{E}(y_s y_t) = \mathbf{E}(y_{\tau(s)} y_{\tau(t)})$$

for any primitive translation τ .

The following theorem shows that this definition of stationarity for processes is consistent with that of stationarity for transfer functions:

Theorem 6 1. The process y is stationary if and only if

$$\mathbf{E}(y_s y_t) = r[d(s, s \wedge t), d(t, s \wedge t)]$$

where $s \wedge t$ is defined in (2.12).

2. If the process u and the transfer function S are both stationary, so is the process Su .

Note that the second statement is an immediate consequence of the first one. More generally, x and y are said to be *jointly* stationary if we have

$$\mathbf{E}(x_s y_t) = r^{xy}[d(s, s \wedge t), d(t, s \wedge t)] \quad (4.16)$$

We define the *cross-spectrum* of x and y as the following power series:

$$R^{xy} \triangleq \sum_{k,l \geq 0} r^{xy}[k, l] \bar{\gamma}^k \gamma^l$$

where $r^{xy}[k, l]$ is the cross-covariance sequence of x and y , cf. (4.16).

Given a stationary transfer function of the form $S = \sum s_{k,l} \bar{\gamma}^k \gamma^l$ (cf. (4.7)), we set

$$S^* \triangleq \sum s_{l,k} \bar{\gamma}^k \gamma^l$$

Then the following formula yields the cross-spectrum of two stationary processes Su and Tu where S and T are stationary transfer functions and u is a stationary process:

$$R^{(Su)(Tu)} = S R^{uu} T^*$$

This formula generalizes a well-known result of the case of standard stationary time series. Finally, Theorem 6 has the following interesting result as a consequence. Pick a point $t_o \in \mathcal{T}$ and order the words $w \in \{\alpha, \beta\}^{*4}$ of length

⁴the language of the words on the alphabet $\{\alpha, \beta\}$

n according to lexicographic order with priority to α : the corresponding set of nodes $t_\alpha w$ is exactly the left-to-right ordered horocycle “segment” in the figure 2, collect the values $y_{t_\alpha w}$ into a vector Y . Then the covariance matrix Σ_Y of Y has the following recursively defined structure:

$$\begin{aligned}\Sigma(r_0) &= r_0 \\ \Sigma(r_0, \dots, r_m) &= \begin{bmatrix} \Sigma(r_0, \dots, r_{m-1}) & r_m U_{m-1} \\ r_m U_{m-1} & \Sigma(r_0, \dots, r_{m-1}) \end{bmatrix} \\ \Sigma_Y &= \Sigma(r_0, \dots, r_n)\end{aligned}$$

where U_m is a $2^m \times 2^m$ -matrix whose entries are 1. It is then easy to show that *the eigenvectors of Σ_Y are the discrete Haar basis*, cf. [5, 13] for more details.

5 Conclusion

In this paper we developed a system theory on the homogeneous dyadic tree as a possible foundation for a *multiscale system theory*. We have shown that the homogeneous tree possesses strange geometric properties that have the following consequences: the double role played by the classical z -transform, namely 1- encoding transfer function and 2- defining stationarity, is split over two different objects —the shifts to encode transfer functions (these are not isometries), and the translations to define stationarity (these are not easily expressed via shifts)—. We sketched two system theories that emphasized on each of these two different objects. Finally a notion of stationary stochastic processes has been introduced.

The major results of this paper may be stated as follows.

1. There is a unique natural way to encode moves on the homogeneous tree, and the corresponding elementary shifts may be used to define and encode transfer functions and develop an associated system and realization theory.
2. There is a unique natural way to define stationarity for both transfer functions and stochastic processes on the homogeneous tree. Such a notion emphasizes “*stochastic fractalness*”, as lengthily discussed in [6, 5]. Note that isotropic processes analyzed in the latter references are a subclass of the stationary processes presented in this paper.
3. Stationary system theory on the dyadic tree is tightly related to the Haar transform (which is the crudest multiscale analysis technique) as expressed by the involvement of the “Haar smoothing” operator $\gamma = \frac{\alpha+\beta}{2}$ and the fact that the restriction of any stationary process at a given scale possesses a covariance function with the discrete Haar basis as eigenvectors.

These results immediately generalize to homogeneous trees with more than 3 branches originating from each node, for instance, multiscale system theory for images would require an homogeneous tree with 5 branches at each node (1 to the coarser scale, and 4 for the pyramid going to finer scale). Proofs of these results will be presented in a full paper, and further results and developments are in progress.

ACKNOWLEDGEMENTS: the authors are indebted to a sizable group for fruitful discussions and stimulations, namely M. Basseville, K.C. Chou, B. Claus, A. Cohen, J-P Conze, B.C. Levy, Y. Meyer, and O. Zeitouni.

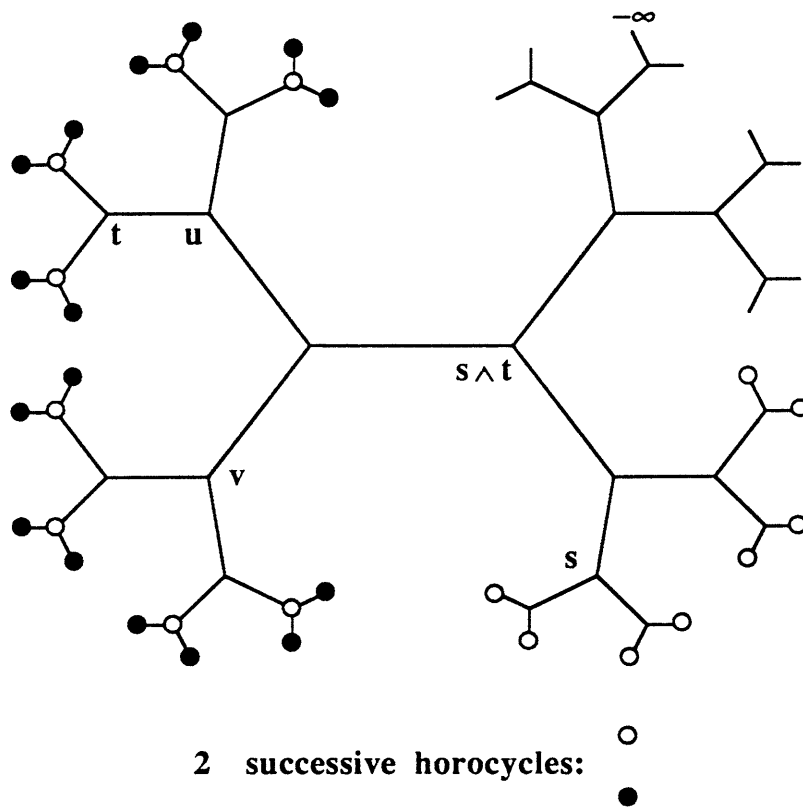


Figure 1: The dyadic homogeneous tree

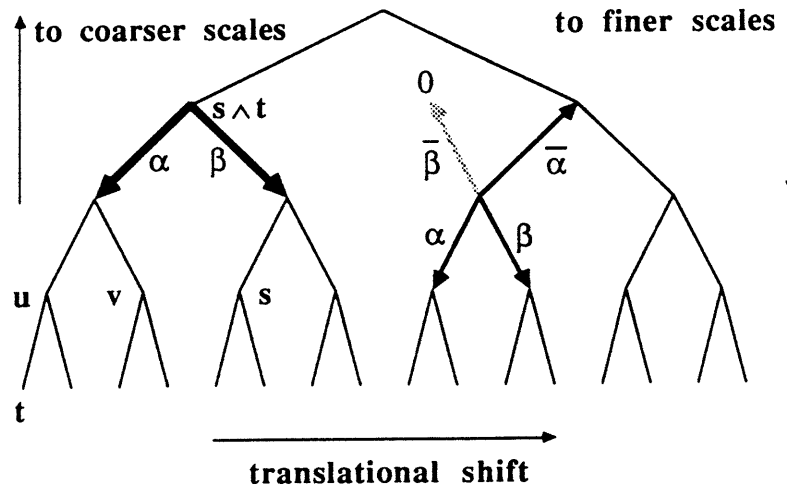


Figure 2: Showing scales and shifts: very thick lines show the moves on the tree, thick lines show the operators on signals (the value at the origin of each arrow is picked at the corresponding end)

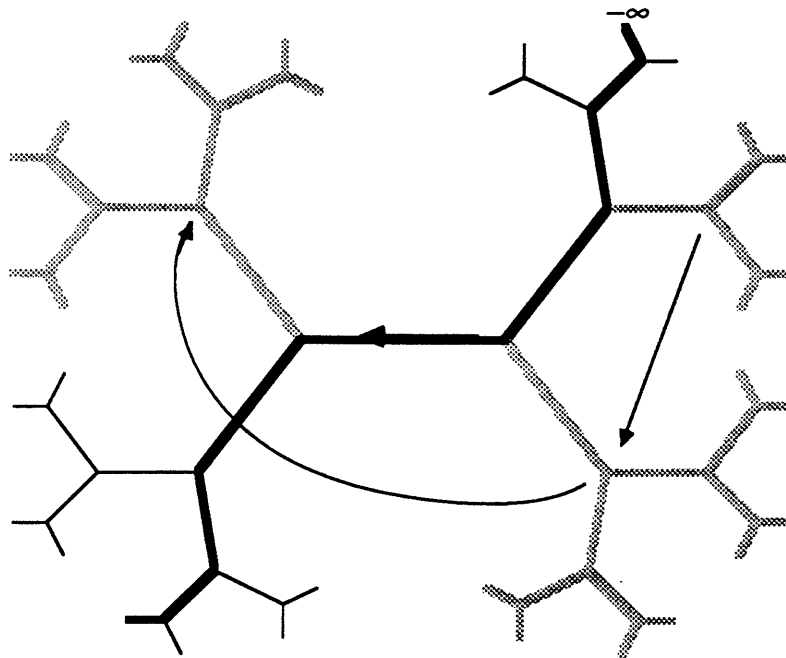


Figure 3: Translations: we show how the $\mathcal{T}_{s_n}^+$ (in grey) are successively mapped

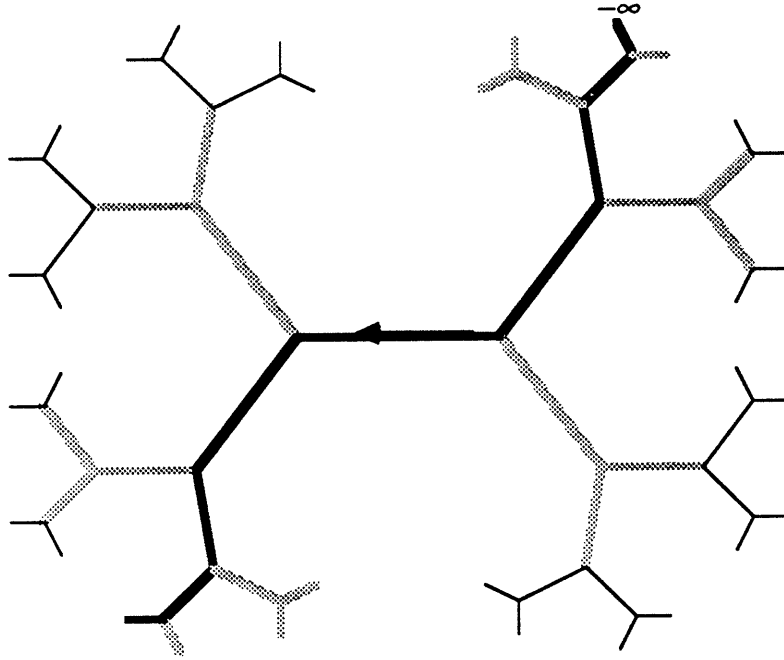


Figure 4: The support of a finite depth transfer function (in grey)

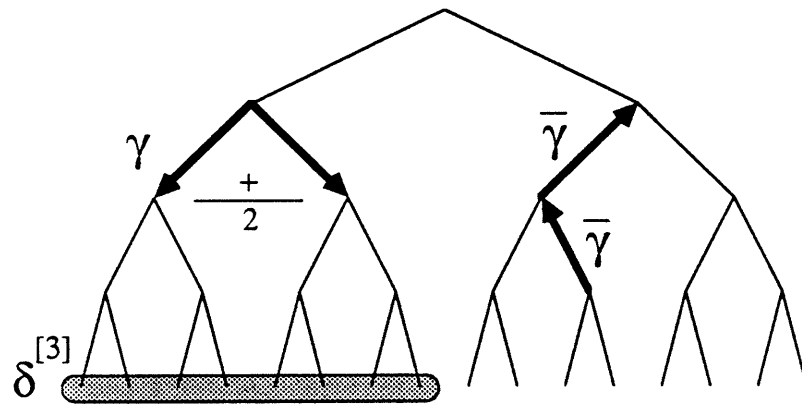


Figure 5: Shifts for stationary transfer functions: the value at the origin of each arrow is picked at the corresponding end and the grey cigare replaces each value by the corresponding average

References

- [1] J.P. ARNAUD, "Fonctions sphériques et fonctions définies positives sur l'arbre homogène," *C.R.A.S.* t. 290, série A (14 Jan. 1980), p. 99–101.
- [2] J.P. ARNAUD, G. LETAC, "La formule de représentation spectrale d'un processus gaussien stationnaire sur un arbre homogène," *Publi. Labo. Stat. and Proba.*, UA 745, Toulouse.
- [3] S. ATTASI, "Systèmes linéaires homogènes à deux indices," *Rapport Laboria*, no. 31, Rocquencourt, 1973.
- [4] M. BARNSLEY, *Fractals Everywhere*, Academic Press, San Diego, 1988.
- [5] M. BASSEVILLE, A. BENVENISTE, AND A.S. WILLSKY, "Multiscale Autoregressive Processes", IRISA and LIDS (MIT) report, submitted for publication.
- [6] M. BASSEVILLE, A. BENVENISTE, K.C. CHOU, AND A.S. WILLSKY, "Multiscale Statistical Signal Processing: Stochastic Processes indexed by Trees" *MTNS 89*, June 19-23, 1989, Amsterdam.
- [7] M. BASSEVILLE AND A. BENVENISTE, "Multiscale statistical signal processing," proc. of the ICASSP-89, 2065-2068, Glasgow, 1989.
- [8] J. BERSTEL, C. REUTENAUER, "Les séries rationnelles et leurs langages," Masson 1984, Collection "Etudes et Recherches en Informatique".
- [9] A. BRANDT, "Multi-level adaptive solutions to boundary value problems," *Math. Comp.*, Vol. 13, pp. 333–390, 1977.
- [10] W. BRIGGS, *A Multigrid Tutorial*, SIAM, Philadelphia, 1987.
- [11] P. CARTIER, "Harmonic analysis on trees," *Proc. Sympos. Pure Math.*, Vol. 26, Amer. Math. Soc., Providence, pp. 419–424, 1974.
- [12] P. CARTIER, "Géométrie et analyse sur les arbres". *Séminaire Bourbaki*, n407, Feb. 1972.
- [13] K.C. CHOU, A.S. WILLSKY, A. BENVENISTE, AND M. BASSEVILLE, "Recursive and Iterative Estimation Algorithms for Multi-Resolution Stochastic Processes," Proc. of the IEEE Conf. on Decision and Control, Dec. 1989.

- [14] K. CHOU, *A Stochastic Modeling Approach to Multiscale Signal Processing*, MIT, Dept. of Electrical Engineering and Computer Science, Ph.D. Thesis, 1990 (in preparation).
- [15] I. DAUBECHIES, A. GROSSMANN, Y. MEYER, "Painless non-orthogonal expansions," *J. Math. Phys.*, vol. 27, p. 1271–1283, 1986.
- [16] I. DAUBECHIES, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Math.*, vol. 91, pp. 909–996, 1988.
- [17] I. DAUBECHIES, "The wavelet transform, time-frequency localization and signal analysis," AT&T Bell Laboratories Report, to appear in *IEEE Trans. on IT*.
- [18] J.L. DUNAU, "Etude d'une classe de marches aléatoires sur l'arbre homogène," in *Arbres homogènes et Couples de Gelfand*, J.P. Arnaud, J.L. Dunau, G. Letac. Publications du Laboratoire de Statistique et Probabilités, Univ. Paul Sabatier, Toulouse, no. 02-83, June 1983.
- [19] B. GIDAS, "A renormalization group approach to image processing problems," *IEEE Trans. on Pattern Anal. and Mach. Int.*, vol. 11, pp. 164–180, 1989.
- [20] J. GOODMAN AND A. SOKAL, "Multi-grid Monte Carlo I. conceptual foundations," Preprint, Dept. Physics New York University, New York, Nov. 1988; to be published.
- [21] P. GOUPILLAUD, A. GROSSMANN, J. MORLET, "Cycle-octave and related transforms in seismic signal analysis," *Geoeexploration*, vol. 23, p. 85–102, 1984/85, Elsevier.
- [22] A. GROSSMANN AND J. MORLET, "Decomposition of Hardy functions into square integrable wavelets of constant shape," *SIAM J. Math. Anal.*, vol. 15, pp. 723–736, 1984.
- [23] A. GROSSMANN, R. KRONLAND-MARTINET, AND J. MORLET, "Reading and Understanding Continuous Wavelet Transforms", in *Wavelets, time-frequency methods and phase space*, 2-20, Proc. of the Int. Conf. on Wavelets, Marseille, Dec 14-18, 1987, J.M. Combes and A. Grossmann Eds., ipti, Springer Verlag, 1989.
- [24] W. HACKBUSCH AND U. TROTTEBERG, eds., *Multigrid Methods*, Springer-Verlag, New York, 1982.

- [25] R. KRONLAND-MARTINET, J. MORLET, A. GROSSMANN, "Analysis of sound patterns through wavelet transforms," *Intl. J. of Pattern Recognition and Artificial Intelligence*, Special Issue on Expert Systems and Pattern Analysis, 1987.
- [26] T. LIN, M. KAWAMATA AND T. HIGUCHI, "New necessary and sufficient conditions for local controllability and observability of 2-D separable denominator systems," *IEEE Trans. Automat. Control*, AC-32, pp. 254-256, 1987.
- [27] S.G. MALLAT, "A compact multiresolution representation: the wavelet model," Dept. of Computer and Info. Science—U. of Penn., MS-CIS-87-69, GRASP LAB 113, Aug. 1987.
- [28] S.G. MALLAT, "A theory for multiresolution signal decomposition: the wavelet representation," Dept. of Computer and Info. Science—U. of Penn., MS-CIS-87-22, GRASP LAB 103, May 1987.
- [29] S.G. MALLAT, "Multiresolution approximation and wavelets," Dept. of Computer and Info. Science—U. of Penn., MS-CIS-87-87, GRASP LAB 80, Sept. 1987.
- [30] B. MANDELBROT, *The Fractal Geometry of Nature*, Freeman, New York, 1982.
- [31] S. MCCORMICK, *Multigrid Methods*, Vol. 3 of the SIAM Frontiers Series, SIAM, Philadelphia, 1987.
- [32] Y. MEYER, *Wavelets and operators*, Proceedings of the Special year in modern Analysis, Urbana 1986/87, published by Cambridge University Press, 1989.
- [33] D. PADDON AND H. HOLSTEIN, eds., *Multigrid Methods for Integral and Differential Equations*, Clarendon Press, Oxford, England, 1985.
- [34] M.J. SMITH AND T.P. BARNWELL, "Exact reconstruction techniques for tree-structured subband coders," *IEEE Trans. on ASSP*, vol. 34, pp. 434-441, 1986.